

Ch 3, Data Analytics: Drawing Conclusions, Part 2

File naming strategies, p 144-146

1. What is the purpose of using effective file naming strategies?

File naming strategies are immediately useful for your solution now, you will also find them handy for years to come, particularly for future studies and works, because it is much easier to be organised when your files are organised too. The main goal is to use a consistent, sensible system.

Formats

2. What is meant by file naming formats?

A file naming format is something that must be obeyed. File naming formats for Windows forbids certain punctuation marks and special characters, prevent the use of names reserved by the OS and ensure that the file name length cannot be exceeded.

Conventions

3. What is the golden rule for conventions?

The golden rule for conventions is that your file names should be short, meaningful and consistent.

4. Describe the following conventions:

a. Characters

With the conventions relating to characters it is important to use numerals for number instead of words (this being because words will not sort numerically and numerals will make the file names shorter), leading zeroes before numerals (files will sort correctly in all OSs, file names without leading zeroes in front of single-digit numbers may otherwise sort into a useless order), plain alpha-numeric characters (most punctuation and other special characters are forbidden by Windows which makes them incompatible for cross-platform compatibility), CamelCase or dashes instead of spaces or underscores in file names (some programs interpret spaces as the end of the file name which will lead to a 404 (file not found) error. Also, using underscores is not recommended because linked on a webpage the underscore will be hidden by the URL) and also use short names (this being that longer names are hard to read and type accurately and will be cut off when viewed in a file manager.)

b. Language

When using the convention of language, it is important to use the Australian English.

c. Case

With the conventions relating to case it is important to use CamelCase (in Characters above) and also to use capitalisation carefully; that being case-sensitive links (this is because most web servers use Linux (which has case sensitive file names); causing your links to a file you have uploaded not working unless you make them case sensitive).

d. Dates

When using the conventions relating to date, the best format is YYYY-MM-DD, which has leading zeroes when necessary.

e. Versions

With the conventions of versions, it is important to use Version numbers (which makes it easier to revert to an earlier saved copy if the current version has become corrupted)

f. Word choice and order

When using the convention relating to word choice and order it is important to use 'Name, The' and Name, A[n]' structure for names that begin with articles [a, an and the,] (this is because sorting files using their first significant word forces a logical sort order), names that omit prepositions and conjunctions if needed (which keeps names short and prevents more important data from being obscured), having file names that start with category names and out more significant data before less significant data, having a 'Family Name, Given name' structure to refer to people (this forces a logical sort order and conforms with standard) and maintain a consistent vocabulary: 'folder' and 'disk' or 'directory' and disk' (as this makes it very easy to avoid ambiguity).

Organising and storing data, p 146-152

1. Why is it important to organise and store your data in an organised way?

If you are disorganised, it will take longer to find the files you want to use. Eventually, mishaps will start to occur, such as losing one important file. Mishaps may then spiral into disasters, such as accidentally deleting directories of documents you have failed to back up.

Hierarchies

2. Explain how a hierarchical structure operates.

Hierarchies make it easy to find a logical place to store and retrieve data in very large collections. Hierarchies are used in site maps, data structure charts and file systems. The structure of a hierarchy is rarely considered strictly right or wrong.

The Cloud and file sharing

3. What are the benefits of using the cloud to store your data?

Some of the benefits of using the cloud to store data is that it is available at any time/anywhere across the globe as long as the user has an internet connection, it is also easy for groups of people to access the data at the same time and make changes and the information is not stored physically which will reduce risks of the loss of data which could occur with an environmental disaster (fire, flood).

4. What is the SPOT strategy and why is this technique popular?

The technique SPOT (which is a single master copy of the document/database on the Cloud) is very popular because it removes the need to organise and track multiple versions of documents in different locations and at different stages of development, permits multiple users to work collaboratively on a document at the same time in different locations and see each other's changes in real time, permits users to 'roll back' a document to its previous state to undo unwanted changes and it gives users the confidence that the document they are using is the definitive version.

Metadata, p 150

5. Explain the term metadata by using an example.

Metadata is descriptive data saved inside a digital data file that can be read by media players and editors (it is basically data about data). An example of this is that metadata about an MP3 will include the track number, artist, cover art, track length, album, the year of recording and so on.

Archiving

6. What does archiving refer to?

Archiving refers to when data is taken offline by moving or copying them to external storage, compressing them to save space and then deleting the original file.

7. What needs to be considered when archiving files?

Files which are stored on offline storage media (such as DVDs, USB flash drives and external HDDs) are inaccessible until attached to a computer and loaded. However, files which are stored on online storage, such as internal HDDs, networks and the cloud, can be loaded immediately.

Synchronising files

8. What are the advantages of data synchronisation?

Utilities designed to synchronise files compare the contents of folders, determine which files are the same or different, newer or older, and copy files back and forth to bring both folders up to date, preserving the timeliness and the integrity of the data.